JAMES N. CHRISTMAN
DIRECT DIAL: 804-788-368
EMAIL: jchristman@hunton.com

June 13, 2002

FILE NO: 48681.000015

Via E-Mail and Facsimile

Ms. Evangeline Tsibris Cummings
Environmental Protection Agency
Office of Environmental Information
Mail Code 2842T
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Attn: Docket ID No. OEI-10014

Dear Ms Cummings:

Accompanying this letter you will find the comments of the Inter-Industry Analytical Group (IIAG) on EPA's Draft Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency.

Yours very truly,

James N. Christman

Enclosure

June 13, 2002

Comments of Inter-Industry Analytical Group (IIAG)
on EPA's Proposed Data Quality Guidelines
(Docket ID No. OEI-10014)

EPA has asked for comment on its "Draft Guidelines for Ensuring And Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency." 67 Fed. Reg. 21,234 (April 30, 2002). These guidelines are required by § 515(a) of the Treasury and General Government Appropriations Act for fiscal year 2001 and respond to an Office of Management and Budget (OMB) guideline directing all federal agencies to develop and implement their own quality guidelines by October 1, 2002. 67 Fed. Reg. 8541 (February 22, 2002).

These are the comments of the Inter-Industry Analytical Group (IIAG). For the purpose of these comments, the IIAG is composed of Alcoa Inc., Alliance of Automobile Manufacturers, American Forest and Paper Association, American Petroleum Institute, Kennecott Utah Copper, and Utility Water Act Group. A public meeting was held in Washington, D.C., on May 15, 2002, to discuss the guidelines. The IIAG appeared and presented comments at that meeting.

IIAG's comments on the draft guidelines are as follows.

1. To satisfy the "objectivity" standard, EPA should adopt the following principles:

a. Data will be disseminated only if accompanied by information on the precision and bias and on the QA/QC standards to which the data have been held.

EPA defines "objectivity" as "whether the disseminated information is being presented in an accurate, clear, complete, and unbiased manner, and as a matter of substance, is accurate, reliable, and unbiased." To ensure objectivity, EPA's guidelines should require that whenever the Agency disseminates data, it will do the following:

1. Specify the test method that was used, confirm that it has been approved officially by EPA, and, where appropriate (as for analytical methods in 40 C.F.R. Part 136), state that it has been promulgated into regulations. If the test method does not meet these requirements, EPA should explain why it believes it can justify disseminating the data despite the use of an unapproved test method.

2. Present all measured data with uncertainty information (precision, bias, *etc.*) so that the public and decisionmakers will be aware of the uncertainty in the data.

3. Specify that appropriate quality assurance/quality control (QA/QC) procedures have been followed and state that the QA/QC results meet the data quality objectives or, if not, explain why EPA believes it can justify disseminating the data despite their failure to meet the objectives.

4. Specify that all values below the level of detection or level of quantification or both should not be used except where they have probative value. Where they do have probative value, EPA should assign the results a value justified or required by the particular application. For example, a measurement below the quantification or detection level should be given a value of zero when averaging data to determine whether measurements are below a numerical standard. On the other hand, a measurement should be taken to be equal to the level of quantification for calculating effluent limitations guidelines.

b. All disseminated data will be "reproducible."

Data used by EPA should be objective. "Objectivity" includes the concept of "reproducibility." OMB defined "reproducibility" as follows:

> With respect to analytic results, "capable of being substantially reproduced" means that independent analysis of the original or supporting data using identical methods would generate similar analytic results, subject to an acceptable degree of imprecision or error.

67 Fed. Reg. 8460 col. 3 (Feb. 22, 2002).

The proposed EPA guidelines do not contain a similar definition. The preamble states instead that "[i]f sufficient transparency is achieved on each of these [four] factors, then an analytic result should meet the 'capable of being substantially reproduced' standard." The four factors include: (1) source of the data used, (2) various assumptions employed, (3) analytic methods applied, and (4) statistical procedures employed.

IIAG supports using the above factors but contends that EPA should establish two additional principles of reproducibility in its guidance. First, as a prerequisite to disseminating data, EPA should require that the data have been generated by a test method that has been validated using an interlaboratory study. Absent such a study, the variability (*i.e.*, the "degree of imprecision") that one can expect to encounter in testing a sample in different laboratories cannot be reliably predicted. This becomes especially important when dealing with low-level measurements, because laboratories do not have reference standards at such low levels for purposes of calibrating their instruments. They must make up those low-level standards by taking higher-level standards and diluting them. This process introduces a bias, the size of which cannot be predicted without a complete interlaboratory study. According to OMB, EPA's guidelines will need to ensure that the Agency's test results can be reproduced by others, subject to an acceptable degree of imprecision. Absent interlaboratory method validation, EPA would be unable to know the expected degree of imprecision associated with its data, and thus it could not claim such imprecision was "acceptable."

An example of a situation where EPA has not established reproducibility involves monitoring data for methylmercury. Both EPA and USGS are now collecting data using test methods that have not been subject to interlaboratory validation. The EPA method has not been added to 40 C.F.R. Part 136, so there has been no public comment process (other than in the

context of the Georgia TMDLs, where these issues have been raised in public comments, but not addressed by EPA). Nor has EPA announced plans to validate or to include Method 1630 in Part 136 at some future date.

As noted above, the OMB Guidelines say that "[w]ith respect to analytic results, 'capable of being substantially reproduced' means that independent analysis of the original or supporting data using identical methods would generate similar analytic results, subject to an acceptable degree of imprecision or error." 67 Fed. Reg. 8460 col. 3. If a test method has not been validated in an interlaboratory study, EPA cannot meet this test. EPA has begun making TMDL decisions, with potentially very substantial economic consequences, using its inadequately validated methylmercury method. EPA may be able to claim that it knows the uncertainty associated with the resulting data for the single laboratory that performed the testing, but it cannot claim that the test results would be reproducible in any other laboratory.

Our concern is not limited to methylmercury. EPA also has been engaging in similar questionable practices with Method 1668 for PCBs, and others. EPA can use the data quality guidelines to ensure tighter controls over the integrity of its data.

EPA itself has recognized the nonreproductibility of method deletion limits ("MDLs") in its recently issued "Analytical Feasibility Support Document for the Six-Year Review of Existing National Primary Drinking Water Regulations" (Draft Report):

> Due to normal day-to-day and run-to-run analytical variability, MDLs may not be reproducible within a laboratory or between laboratories.

Because the MDL is not reproducible, quantitation values derived as a multiplier of the MDL (*e.g.*, the minimum level or "ML") also are not reproducible. EPA should clarify that neither the MDL nor values derived based on the MDL should be used as the basis for determining what data are suitable to be disseminated. To the extent EPA may claim that it currently has no alternative concepts to use for determining which trace level measurement data can be disseminated, it should move aggressively to develop such concepts. The Agency is currently engaged in a rulemaking to reassess the MDL and ML. It should take "reproducibility" into account as a criterion in deciding whether those concepts are acceptable, or whether alternative concepts (concepts that allow reproducibility, or a least provide a basis for determining the level of uncertainty) are necessary.

Second, EPA should establish, as a prerequisite to disseminating data, that the data must be generated by a test method that contains adequate and mandatory QA/QC provisions. Mandatory QA/QC provisions ensure that laboratories using the test will run it properly and thereby achieve essentially the same results as any other laboratory using the method. Without these QA/QC requirements, EPA cannot assert that its data are "reproducible." The data arising from EPA or EPA contract laboratories using test methods lacking mandatory QA/QC provisions would likely be more variable than the test method was expected to perform when it was originally validated. That is because validation studies typically specify very rigid QA/QC requirements. Again, without QA/QC requirements, EPA would be unable to know the expected degree of imprecision associated with its data. Such data are not reproducible.

The best example is EPA's whole effluent toxicity ("WET") methods. We refer EPA to the WET Coalition comments on the WET Test Methods proposed rule. 66 Fed. Reg. 49,794 (Sept. 28, 2001). Many of the QA/QC provisions in that rule are not mandatory. Many only *appear* to be mandatory; they use the words "must" or "shall" but later state that the regulator can deem the test result acceptable even if the laboratory deviated from the so-called requirement. Yet the rule does not include any criteria by which the regulator is supposed to make these judgments. So where EPA is the data-gathering entity and its WET laboratories fail to follow all the so-called requirements, if EPA simply excuses those deviations, as the test protocols allow, how can it say that the data are "reproducible"?

2. Public comment is no substitute for a data correction process.

In § 5.4 of the proposed guidelines, EPA states that it will consider requests for correction unless the request "pertains to EPA actions, where a mechanism by which to submit comments to the Agency is already provided." That broad exemption is entirely inconsistent with the basic objectives of the statute creating the data quality guidelines process and with OMB's guidelines to the agencies. For one thing, the ordinary public comment process for proposed rules is not sufficiently focused on data quality to provide an adequate means of correcting errors. If EPA does not correct data quality problems before it issues a final rule, and also does not provide a right to an administrative appeal of the Agency's judgment to rely on certain data to support its final decision, then it will forfeit the opportunity to avoid the inefficiency of having to defend in court a final decision that was developed on the basis of faulty data. In addition, sometimes there is a long time between publication of the draft rule and promulgation of the final rule. Data problems should be addressed early in the rulemaking process rather than late.

3. EPA should not exempt contractors, grantees, or States.

EPA's exemptions from what constitute "dissemination" of data are too broad. In the first place, EPA proposes to exempt data given to its contractors and grantees. But the work of contractors and grantees is often used to support agency initiatives. Data to be disseminated to contractors and grantees should not be exempt from the data quality guidelines, especially when EPA expects that those data will be used to develop a product (*e.g.*, a model) that will support EPA actions that will be subject to the guidelines.

In the second place, EPA proposes (at line 474 of the proposed guidelines) to exempt States. In effect, EPA proposes to exclude from the definition of "dissemination" (and thus from data review) any information it generates and gives to States to use. Because regulatory authority is extensively delegated to States under the Clean Water Act, Clean Air Act, and other environmental programs, it would be intolerable to exempt from review data given to States. EPA ultimately must review and approve or disapprove many of those State decisions. Given that the data quality guidelines would apply to those EPA decisions, and that EPA approval of a State decision constitutes EPA approval of the data on which the State relied, EPA's decision, and the underlying data, would be subject to the data quality guidelines. It would be far more efficient to subject the data to be given the State to the data quality guidelines, and to have the opportunity to correct data problems, before the State relies on those data to perform its delegated tasks.

4. EPA must develop a process for consistently applying its QA/QC procedures.

EPA has developed several useful documents addressing data quality. The following guidance and requirements are examples:

- *Guidance for Data Quality Assessment* (July 2000);
- *Guidance for the Data Quality Objectives Process* (August 2000);
- *Guidance for Preparing Standard Operating Procedures (SOPs)* (March 2001);
- *EPA Requirements for Quality Management Plans* (March 2001); and
- *EPA Requirements for Quality Assurance Project Plans* (March 2001).

Unfortunately, EPA applies these documents inconsistently. As required by the OMB guidelines, EPA needs to develop a consistent "process" for implementing its QA/QC guidelines. It is not enough simply to continue the present *ad hoc* application of these guidelines. In the method validation study for the WET rulemaking, for example, EPA determined that its contract laboratories deviated from the strict QA/QC procedures that it had required in the validation study plan. EPA chose to overlook those deviations and decided to rely on the questionable data. See Comments of the WET Coalition on EPA's Proposal to Ratify or Withdraw WET Test Methods (Jan 11, 2002) (available in the WET rule docket or upon request).

The Agency itself acknowledges deficiencies in implementing its data quality guidance. *See* "Lessens Learned About Designing, Developing, and Disseminating Environmental Information Products," Nov. 17, 2000, at p. 23 (available online at *www.epa.gov/oei/pdf/OIAA_Lessons-learned.pdf*). EPA should establish a specific "process," as OMB requires.

5. Trade associations should be entitled to review of data.

EPA's proposed guidelines allow data correction requests to be submitted only by an "affected person." See Proposed Guidelines §5.2, line 715. This means a person who may benefit or be harmed by the disseminated information.

The review procedures need to be clarified to include trade associations and *ad hoc* organizations with affected members. This interpretation is applied in the federal courts and should apply to the data quality guidelines as well. EPA should consider the efficiency such an interpretation would provide, particularly if more than one affected person decide to seek correction of the same data. Unless a trade association or similar organization can seek correction on behalf of its individual members, EPA will have to deal with the administrative burdens of multiple requests.

6. EPA should clarify that other agencies are allowed to invoke EPA's data correction process.

EPA should make clear that other agencies, both State and federal, may ask for data correction, as well as natural persons, trade associations, and *ad hoc* organizations. There is no legitimate reason for excluding government agencies from the process. Their early input can avoid disputes that may be difficult to resolve once the allegedly flawed data are incorporated into models or other analyses.

7. Appeal of data correction decisions should go to the Chief Information Officer.

EPA's proposed appeal process calls for the responsible Assistant Administrator or Regional Administrator to make the final decision on appeal, with advice from a panel. This process will not be effective.

The appeal should go instead to the Chief Information Officer (CIO), who should have authority to call experts for assistance in understanding complex technical matters. Neither the AA nor the RA is in a position to render an unbiased decision in an appeal, given that they initially were responsible for approving the decision based on the allegedly flawed data. The CIO does not have that connection.

For the appeal process to be equitable and effective, the EPA Administrator must convey to EPA staff the importance and benefit of the data quality requirements to the entire agency. Further, the CIO must receive clear and unambiguous authority from the Administrator to review and rule on the appeal using whatever resources are necessary to reach a decision.

8. A deadline needs to be set for EPA responses to requests for correction.

EPA needs to set a time limit by which it will respond to requests for correction of data. IIAG suggests 30 days from receipt of a request for data correction and 60 days from receipt of an appeal of the Information Owner's decision, with perhaps some exceptions for good cause.

9. OMB's guidelines are mandatory.

EPA says (at lines 401-06) that the guidelines do not impose any legally binding requirements or obligations on EPA. EPA says it retains discretion to violate the guidelines on a case-by-case basis "where appropriate."

Despite EPA's assertion to the contrary, the OMB guidelines are a legal requirement. In House Report 105-592 (pp. 49-50) for the 1999 Appropriations Act (H.R. 4104) there was a directive to OMB for the development of rules "ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by Federal agencies." The report said that the committee "urges" the Office of Management and Budget to issue quality guidelines. About a year later, based in part on the earlier language, Congress included specific data quality language in § 515 of the Treasury and General Government Appropriations Act for fiscal year 2001. This time Congress said that the Director of OMB "shall" issue guidelines. See W. Olsen, "The Federal Data Quality Act" (2002).

Moreover § 515(a) says that the guidelines are to be issued "under sections 3504(d)(1) and 3516" of 44 U.S.C. that provide policy and procedural guidance to federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information. Section 3516 says that the Director shall promulgate "rules, regulations, or procedures" necessary to exercise the authority provided by the subchapter (meaning subchapter I, 44 U.S.C. §§ 3501-3520, addressing "Federal Information Policy").

Moreover, part of the subchapter is § 3506(a)(1), which says that the head of each agency shall be responsible for complying with the requirements of the subchapter and related policies established by the Director. Under § 3506(a)(3) an agency's Chief Information Officer heads an office responsible for ensuring agency compliance with and prompt, efficient, and effective implementation of the information policies and information resources management responsibilities established under the chapter (44 U.S.C. §§ 3501 *et seq.*). Finally, § 3506(b) says that each agency "shall" manage information resources to improve the integrity, quality, and utility of information to all users.

In short, Congress intended that the data quality guidelines be followed, and not just at the agencies' discretion.

10. EPA needs to define original and supporting data that it says may not be subject to the high degree of transparency required of analytical results.

In § 3.3 of the proposed guidelines, starting at line 640, EPA says that "[o]riginal and supporting data may not be subject to the high and specific degree of transparency required of analytic results." This exception is troubling, because it seems to open up a wide exception for "original and supporting data." EPA needs to define exactly what it means by "original and supporting data" in this provision.

11. EPA needs to clarify the scope of the term "influential."

Whether information is "influential" determines whether it is subject to the data quality process. EPA needs to clarify the scope of the term "influential."

The OMB Guidelines define "influential" to mean that "the agency can reasonably determine that dissemination of the information will have or does have a clear and substantial impact on important public policies or important private sector decisions." 67 Fed. Reg. 8460 col. 3 (February 22, 2002). EPA's proposed guidelines say that information disseminated in support of "top agency actions" is what is "influential." But some very significant decisions are signed by an Assistant Administrator rather than by the Administrator. Such actions should be defined as "top actions."

Likewise, many significant guidance documents or other decisions are signed by Division Directors or even program staff. These too could have a very substantial impact. Finally, decisions made at the Regional level, such as approval of total maximum daily loads that may be based on flawed data, also should be subject to the process. In some cases, program staff, rather than the Regional Administrator, will make extremely important decisions (*e.g.*, sending a letter approving a model the State has developed as part of a TMDL that will impose multi-million

dollar wasteload reduction obligations on dischargers). Those decisions, even though not made by top agency officials, have the same effect, and thus must be subject to the guidelines.

12. Data that are not generated by EPA but that are later endorsed by and disseminated by EPA in a decision should be subject to the same quality standards that EPA applies to itself.

Sometimes EPA will disseminate data that it has not generated. A State may make a decision, based on data, and EPA will then make a decision to approve or disapprove the State action, relying on the State's data. For example, when a State develops a TMDL, EPA is responsible for reviewing the TMDL and disseminating its decision approving or disapproving it (based on the underlying data). State and other third party-generated data, when EPA believes there is a reasonable chance it later will adopt the data as a basis for its own decision, should be held to the same standards that EPA applies to its own work.

Also, as part of the pre-dissemination review process, EPA needs to have an outreach program to let the States know in advance EPA's expectations for data quality. For example, the Delaware River Basin Commission ("DRBC") currently is developing a TMDL for PCBs in the Delaware Estuary. Thus far, all sampling in support of the TMDL development process has involved test methods that EPA has neither validated nor otherwise approved. The DRBC has not independently validated the test method either. Given that EPA Region III is ultimately responsible for reviewing the TMDL and deciding whether to approve it (along with the data on which the TMDL is based), Region III should inform the DRBC early in the process how it intends to review data generated with unapproved test methods. That standard should be no different from the one EPA applies to data it has generated directly. (See our comment, above, on what that standard should be.) If EPA sets a standard of quality for its own data, it should not compromise that standard when reviewing the data supporting a State or other decision that EPA may endorse.

13. IIAG requests the opportunity to review and offer comments on the next revision of the guidelines.

EPA's draft guidelines provide a useful discussion of many of the key issues, but they offer few details on the Agency's proposed approach. IIAG and other reviewers, therefore, have not had the opportunity to comment directly on the guidelines the Agency proposes to submit to OMB.

EPA has expressed an intent to substantially revise its guidelines. IIAG requests the opportunity to review and offer comments on the next revision, whether or not that opportunity comes after August 1, when EPA is supposed to have its guidelines to OMB.

Respectfully Submitted,

Steven J. Koorse
James N. Christman
Counsel to IIAG

cc: Ms. Nancy W. Wentworth (via e-mail)
Ms. Elaine G. Stanley (via e-mail)
Mr. Mark A. Luttner (via e-mail)
Mr. Jeffrey C. Worthington (via e-mail)
Ms. Brooke J. Dickson (via e-mail)